Locally Equivalent Weights for Bayesian MrP

Ryan Giordano, Alice Cima, Erin Hartman, Jared Murray, Avi Feller

UT Austin Statistics Seminar September 2025

Are US non-voters becoming more Republican?

Blue Rose research says yes:

“Politically disengaged voters have become much more Republican, and because less-engaged voters swung away from [Democrats], an expanded electorate meant a more Republican electorate.”

(Blue Rose Research 2024)
(major professional pollsters)

***On Data and Democracy* says no:**

“Claims of a decisive pro-Republican shift among the overall non-voting population are not supported by the most reliable, large-scale post-election data currently available.”

(Bonica et al. 2025)
(major professional researchers)

Are US non-voters becoming more Republican?

Blue Rose research says yes:

“Politically disengaged voters have become much more Republican, and because less-engaged voters swung away from [Democrats], an expanded electorate meant a more Republican electorate.”

(Blue Rose Research 2024)
(major professional pollsters)

***On Data and Democracy* says no:**

“Claims of a decisive pro-Republican shift among the overall non-voting population are not supported by the most reliable, large-scale post-election data currently available.”

(Bonica et al. 2025)
(major professional researchers)

- The problem is very hard (it’s difficult to accurately poll non–voters)
- Different data sources
- ⋆⋆⋆ **Different statistical methods**
  - Blue Rose uses Bayesian hierarchical modeling (MrP)
  - On Data and Democracy is using calibration weighting (CW)

Are US non-voters becoming more Republican?

Blue Rose research says yes:

“Politically disengaged voters have become much more Republican, and because less-engaged voters swung away from [Democrats], an expanded electorate meant a more Republican electorate.”

(Blue Rose Research 2024)
(major professional pollsters)

***On Data and Democracy* says no:**

“Claims of a decisive pro-Republican shift among the overall non-voting population are not supported by the most reliable, large-scale post-election data currently available.”

(Bonica et al. 2025)
(major professional researchers)

- The problem is very hard (it’s difficult to accurately poll non–voters)
- Different data sources
- ⋆⋆⋆ **Different statistical methods**
  - Blue Rose uses Bayesian hierarchical modeling (MrP)
  - On Data and Democracy is using calibration weighting (CW)

Our contribution

We define “MrP local equivalent weights” (MrPlew) that:

- Are easily computable from MCMC draws and standard software, and
- Provide MrP versions of key diagnostics that motivate calibration weighting.

$\Rightarrow$ **MrPlew provides direct comparisons between MrP and calibration weighting.**

Outline

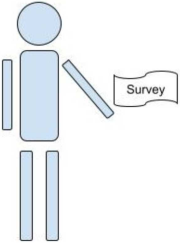
- Introduce the statisical problem and two methods (CW and MrP)
- Describe covariate balance, one of the classical CW diagnostics
- Define MrPlew weights and connect them to covariate balance
- Example of real-world results
- Future directions

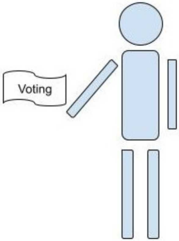
The basic problem

We have a survey population, for whom we observe:

- Covariates $\mathbf{x}$ (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to "do you support Trump")

We want the average response in a target population, in which we observe only covariates.





Observe $(\mathbf{x}_i, y_i)$ for $i = 1, \ldots, N_S$

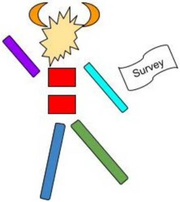
Observe $\mathbf{x}_j$ for $j = 1, \ldots, N_T$

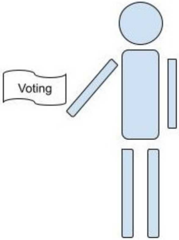
The basic problem

We have a survey population, for whom we observe:

- Covariates $\mathbf{x}$ (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to “do you support Trump”)

We want the average response in a target population, in which we observe only covariates.





Observe $(\mathbf{x}_i, y_i)$ for $i = 1, \ldots, N_S$ | Observe $\mathbf{x}_j$ for $j = 1, \ldots, N_T$

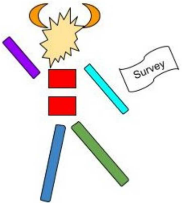
The problem is that the populations may be very different.

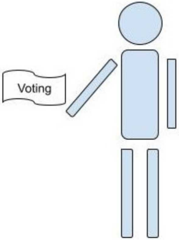
The basic problem

We have a survey population, for whom we observe:

- Covariates **x** (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to “do you support Trump”)

We want the average response in a target population, in which we observe only covariates.





Observe $(\mathbf{x}_i, y_i)$ for $i = 1, \ldots, N_S$ | Observe $\mathbf{x}_j$ for $j = 1, \ldots, N_T$

The problem is that the populations may be very different.

Our survey results may be biased.

How can we use the covariates to say something about the target responses?

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target population y_j.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target population y_j.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

▶ Choose "calibration weights" w_i
using only the regressors $\mathbf{x}$
(e.g. raking weights)

Bayesian hierarchical modeling (MrP)

▶ Choose $\mathbb{E}\left[y|\mathbf{x},\theta\right] = m(\theta^\intercal \mathbf{x})$,
choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$
(e.g. Hierarchical logistic regression)

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target population y_j.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

- ▶ Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)
- ▶ Take $\hat{\mu}_{\mathrm{CW}} = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i y_i$

Bayesian hierarchical modeling (MrP)

- ▶ Choose $\mathbb{E}\left[y|\mathbf{x},\theta\right] = m(\theta^\intercal \mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)
- ▶ Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}_{\mathrm{MrP}} = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{y}_j$

Two approaches

We want $\mu := \frac{1}{N_T}\sum_{j=1}^{N_T} y_j$, but don't observe target population y_j.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

- Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)
- Take $\hat{\mu}_{\text{CW}} = \frac{1}{N_S}\sum_{i=1}^{N_S} w_i y_i$
- Dependence on y_i is clear

Bayesian hierarchical modeling (MrP)

- Choose $\mathbb{E}\left[y|\mathbf{x},\theta\right] = m(\theta^{\intercal}\mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)
- Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}_{\text{MrP}} = \frac{1}{N_T}\sum_{j=1}^{N_T}\hat{y}_j$
- Dependence on y_i very complicated (Typically via MCMC draws from $\mathcal{P}(\theta|\text{Survey data})$)

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target population y_j.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

- ▶ Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)
- ▶ Take $\hat{\mu}_{\mathrm{CW}} = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i y_i$
- ▶ Dependence on y_i is clear
- ▶ Weights give interpretable diagnostics:
  - Frequentist variability
  - Partial pooling
  - Regressor balance

Bayesian hierarchical modeling (MrP)

- ▶ Choose $\mathbb{E}\left[y|\mathbf{x},\theta\right] = m(\theta^\intercal \mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)
- ▶ Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}_{\mathrm{MrP}} = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{y}_j$
- ▶ Dependence on y_i very complicated (Typically via MCMC draws from $\mathcal{P}(\theta|\text{Survey data})$)
- ▶ **Black box**

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{j=1}^{N_T} y_j$, but don't observe target population y_j.

- Assume $p(y|\mathbf{x})$ is the same in both populations,
- But the distribution of $\mathbf{x}$ may be different in the survey and target.

Calibration weighting (CW)

- ▶ Choose "calibration weights" w_i using only the regressors $\mathbf{x}$ (e.g. raking weights)
- ▶ Take $\hat{\mu}_{\mathrm{CW}} = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i y_i$
- ▶ Dependence on y_i is clear
- ▶ Weights give interpretable diagnostics:
  - Frequentist variability
  - Partial pooling
  - Regressor balance

Bayesian hierarchical modeling (MrP)

- ▶ Choose $\mathbb{E}\left[y|\mathbf{x},\theta\right] = m(\theta^{\intercal}\mathbf{x})$, choose prior $\mathcal{P}(\theta|\Sigma)\mathcal{P}(\Sigma)$ (e.g. Hierarchical logistic regression)
- ▶ Take $\hat{y}_j = \mathbb{E}_{\mathcal{P}(\theta|\text{Survey data})}\left[y|\mathbf{x}_j\right]$ and $\hat{\mu}_{\mathrm{MrP}} = \frac{1}{N_T} \sum_{j=1}^{N_T} \hat{y}_j$
- ▶ Dependence on y_i very complicated (Typically via MCMC draws from $\mathcal{P}(\theta|\text{Survey data})$)
- ▶ **Black box**

$\leftarrow$ We open this box, providing analogues of all these diagnostics

Prior work

Gelman (2007b) observes that MrP is a CW estimator when one uses linear regression to form $\hat{y}$:

$$\hat{\mu}_{\text{MrP}} = \frac{1}{N_T}\sum_{j=1}^{N_T} \hat{y}_j = \frac{1}{N_T}\sum_{j=1}^{N_T} \underbrace{\mathbf{x}_j^\intercal \hat{\beta}}_{\text{Linear in } y_i}$$

Most existing literature on comparing CW and MrP focus on such linear models. [1]

[1] For example, Gelman (2007b), B., F., and H. (2021), and Chattopadhyay and Zubizarreta (2023).

Prior work

Gelman (2007b) observes that MrP is a CW estimator when one uses linear regression to form $\hat{y}$:

$$\hat{\mu}_{\mathrm{MrP}} = \frac{1}{N_T}\sum_{j=1}^{N_T} \hat{y}_j = \frac{1}{N_T}\sum_{j=1}^{N_T} \underbrace{\mathbf{x}_j^\intercal \hat{\beta}}_{\text{Linear in } y_i}$$

Most existing literature on comparing CW and MrP focus on such linear models. [1]

But what if you use a non–linear link function? Or a hierarchical model?

> *"It would also be desirable to use nonlinear methods ... but then it would seem difficult to construct even approximately equivalent weights. Weighting and fully nonlinear models would seem to be completely incompatible methods." — (Gelman 2007a)*

[1] For example, Gelman (2007b), B., F., and H. (2021), and Chattopadhyay and Zubizarreta (2023).

Equivalent weights for (some) logistic regression MrP

Consider logistic regression MrP:

- Model $m(\mathbf{x}^\intercal \beta) = \text{Logistic}(\mathbf{x}^\intercal \beta)$
- Let $\hat{\beta}$ be the MLE
- MrP is $\hat{\mu}_{\text{MrP}} = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\beta})$.

Suppose $\mathbf{x} \in \mathcal{X}$ is discrete and saturated.

Then logistic MrP is a CW estimator!

- Let $\overline{y}_S^c$ denote the survey average among $\mathbf{x} = c$ for $c \in \mathcal{X}$
- For $\mathbf{x} = c$, $m(\hat{\beta}^\intercal \mathbf{x}) = \overline{y}_S^c$
- Let N_S^c (or N_S^c) denote the # of survey (or target) observations with $\mathbf{x}_n = c$.

$$\hat{\mu}_{\text{MrP}} = \frac{1}{N_T} \sum_{j=1}^{N_T} m(\mathbf{x}_j^\intercal \hat{\beta}) = \frac{1}{N_T} \sum_{c \in \mathcal{X}} \underbrace{N_T^c \overline{y}_S^c}_{\text{Linear in } y_i} = \frac{1}{N_S} \sum_{i=1}^{N_S} w_i^{\text{MrP}} y_i$$

For $w_i^{\text{MrP}} = \dfrac{N_T^c / N_T}{N_S^c / N_S}$ when $\mathbf{x}_i = c$.

The weights can look very different!

Does this mean anything? Are the differences important?

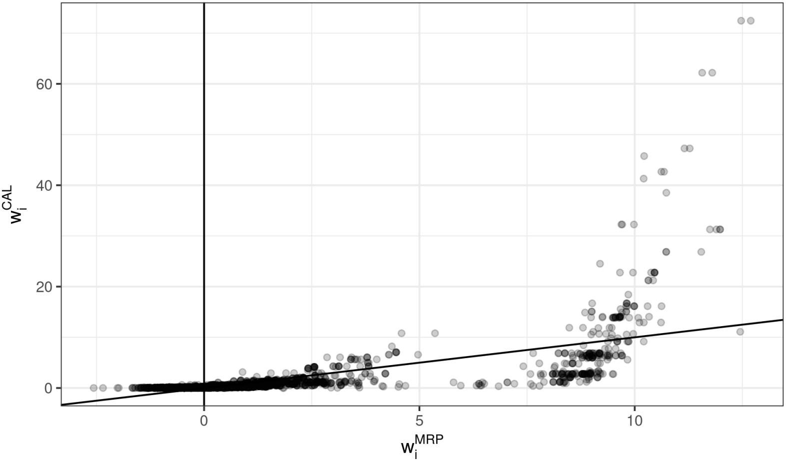


Figure 1: Comparison between raking and MrPlew weights for a particular example

What are we weighting for?[2]

We want:

$$\text{Target average response} = \frac{1}{N_T}\sum_{j=1}^{N_T} y_j \approx \frac{1}{N_S}\sum_{i=1}^{N_S} w_i y_i = \text{Weighted survey average response}$$

We can't check this, because we don't observe y_j.

[2] Pun attributable to Solon, Haider, and Wooldridge (2015)

What are we weighting for?[2]

We want:

$$\text{Target average response} = \frac{1}{N_T}\sum_{j=1}^{N_T} y_j \approx \frac{1}{N_S}\sum_{i=1}^{N_S} w_i y_i = \text{Weighted survey average response}$$

We can't check this, because we don't observe y_j. But we can check whether:

$$\frac{1}{N_T}\sum_{j=1}^{N_T} \mathbf{x}_j = \frac{1}{N_S}\sum_{i=1}^{N_S} w_i \mathbf{x}_i$$

Such weights satisfy "covariate balance" for **x**.

You can check covariate balance for any calibration weighting estimator, and any function $f(\mathbf{x})$.

[2] Pun attributable to Solon, Haider, and Wooldridge (2015)

What are we weighting for?[2]

We want:

$$\text{Target average response} = \frac{1}{N_T}\sum_{j=1}^{N_T} y_j \approx \frac{1}{N_S}\sum_{i=1}^{N_S} w_i y_i = \text{Weighted survey average response}$$

We can't check this, because we don't observe y_j. But we can check whether:

$$\frac{1}{N_T}\sum_{j=1}^{N_T} \mathbf{x}_j = \frac{1}{N_S}\sum_{i=1}^{N_S} w_i \mathbf{x}_i$$

Such weights satisfy "covariate balance" for $\mathbf{x}$.

You can check covariate balance for any calibration weighting estimator, and any function $f(\mathbf{x})$.

Even more, covariate balance is the criterion for a popular class of calibration weight estimators:

Raking calibration weights

"Raking" selects weights that

- Are as "close as possible" to some reference weights
- Under the constraint that they balance some selected regressors.

[2] Pun attributable to Solon, Haider, and Wooldridge (2015)

Generalized covariate balance checks

We want to balance $f(\mathbf{x})$ because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Generalized covariate balance checks

We want to balance $f(\mathbf{x})$ because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Balance–informed sensitivity check (BISC) (informal)

Pick a small $\delta > 0$ and an $f(\cdot)$. Define a *new response variable* $\tilde{y}$ such that

$$\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x}).$$

We know the change this is supposed to induce in the target population.

Covariate balance checks whether our estimators produce the same change.

Generalized covariate balance checks

We want to balance $f(\mathbf{x})$ because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Balance–informed sensitivity check (BISC) (formal)

Pick a small $\delta > 0$ and an $f(\cdot)$. Define a *new response variable* $\tilde{y}$ such that

$$\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x}).$$

We know the expected change this perturbation produces in the target distribution:

$$\mathbb{E}\left[\mu(\tilde{y}) - \mu(y)|\mathbf{x}\right] = \frac{1}{N_T}\sum_{j=1}^{N_T}\left(\mathbb{E}\left[\tilde{y}|\mathbf{x}_p\right] - \mathbb{E}\left[y|\mathbf{x}_p\right]\right) = \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j)$$

Then, check whether your estimator $\hat{\mu}(\cdot)$ produces the same change for observed $\tilde{y}, y$:

$$\underbrace{\hat{\mu}(\tilde{y}) - \hat{\mu}(y)}_{\substack{\text{Replace weighted averages}\\ \text{with changes in an estimator}}} \overset{\text{check}}{\approx} \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j).$$

Generalized covariate balance checks

We want to balance $f(\mathbf{x})$ because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Balance–informed sensitivity check (BISC) (formal)

Pick a small $\delta > 0$ and an $f(\cdot)$. Define a *new response variable* $\tilde{y}$ such that

$$\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x}).$$

We know the expected change this perturbation produces in the target distribution:

$$\mathbb{E}\left[\mu(\tilde{y}) - \mu(y)|\mathbf{x}\right] = \frac{1}{N_T}\sum_{j=1}^{N_T}\left(\mathbb{E}\left[\tilde{y}|\mathbf{x}_p\right] - \mathbb{E}\left[y|\mathbf{x}_p\right]\right) = \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j)$$

Then, check whether your estimator $\hat{\mu}(\cdot)$ produces the same change for observed $\tilde{y}, y$:

$$\underbrace{\hat{\mu}(\tilde{y}) - \hat{\mu}(y)}_{\text{Replace weighted averages with changes in an estimator}} \overset{\text{check}}{\approx} \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j).$$

When $\hat{\mu}(\cdot) = \hat{\mu}_{\mathrm{CW}}(\cdot)$, BISC recovers the standard covariate balance check.

When $\hat{\mu}(\cdot) = \hat{\mu}_{\mathrm{MrP}}(\cdot)$ and δ is small, BISC recovers our proposal.

Generalized covariate balance for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$.

Generalized covariate balance for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$.

Problem: $\hat{\mu}_{\text{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$ may be small.

Generalized covariate balance for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$.

Problem: $\hat{\mu}_{\text{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\text{MrP}} (\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\text{MrP}} := \frac{d}{dy_i} \hat{\mu}_{\text{MrP}}(y).$$

Generalized covariate balance for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$.

Problem: $\hat{\mu}_{\mathrm{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} (\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\mathrm{MrP}} := \frac{d}{dy_i} \hat{\mu}_{\mathrm{MrP}}(y).$$

Only valid **locally**: for "small" $\|\tilde{y} - y\|$.

But that's okay — we are considering small perturbations.

Generalized covariate balance for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$.

Problem: $\hat{\mu}_{\text{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\text{MrP}} (\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\text{MrP}} := \frac{d}{dy_i} \hat{\mu}_{\text{MrP}}(y).$$

Use in BISC: For a wide set of judiciously chosen $f(\cdot)$, check

$$\underbrace{\delta \sum_{i=1}^{N_S} w_i^{\text{MrP}} f(\mathbf{x}_i)}_{\approx \hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)} \overset{\text{check}}{\approx} \delta \frac{1}{N_T} \sum_{j=1}^{N_T} f(\mathbf{x}_j).$$

This a **sensitivity analysis** that formally coincides with a **balance check**.

Generalized covariate balance for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$.

Problem: $\hat{\mu}_{\text{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\text{MrP}}(\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\text{MrP}} := \frac{d}{dy_i}\hat{\mu}_{\text{MrP}}(y).$$

Computation: The weights are given by weighted averages of posterior covariances[3].

They can be easily computed with standard software[4] **without re–running MCMC**.

[3]**G.**, Broderick, and Jordan 2018.
[4]We use `brms` (Bürkner 2017).

Generalized covariate balance for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$.

Problem: $\hat{\mu}_{\mathrm{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} (\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\mathrm{MrP}} := \frac{d}{dy_i} \hat{\mu}_{\mathrm{MrP}}(y).$$

Theory: We state conditions under which, as $\delta \to 0$, and $N \to \infty$,

- The error is of lower order than the MrPlew term,
- *Uniformly* over a very wide class of $f(\cdot)$.

Uniformity is the hard part, but this justifies using MrPlew to *identify* problematic $f(\cdot)$.

Builds on earlier work on uniform error bounds for Bernstein–von Mises theorem(–ish) results[3].

[3] **G.** and Broderick 2024; Kasprzak, **G.**, and Broderick 2025.

Real Data: Marital Name Change Survey

Analysis of changing names after marriage (based on Alexander (2019)).

- **Target population:** ACS survey of US population 2017–2022[4]
- **Survey population:** Marital Name Change Survey[5]
- **Respose:** Did the female partner keep their name after marriage?
- For regressors, use bins of age, education, state, and decade married.

Survey observations:	$N_S = 4,364$
Target observations (rows):	$N_T = 4,085,282$

Uncorrected survey mean:	$\frac{1}{N_S} \sum_{i=1}^{N_S} y_i = 0.462$
Raking:	$\hat{\mu}_{\mathrm{CW}} = 0.263$
MrP:	$\hat{\mu}_{\mathrm{MrP}} = 0.288 \quad (\text{Post. sd} = 0.0169)$

[4]Ruggles et al. 2024.

[5]Cohen 2019.

Covariate balance for primary effects

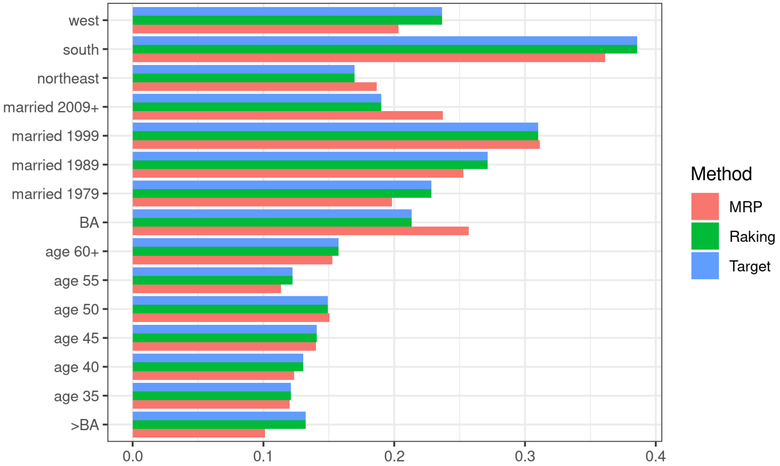


Figure 2: Imbalance plot for primary effects

Covariate balance for interaction effects

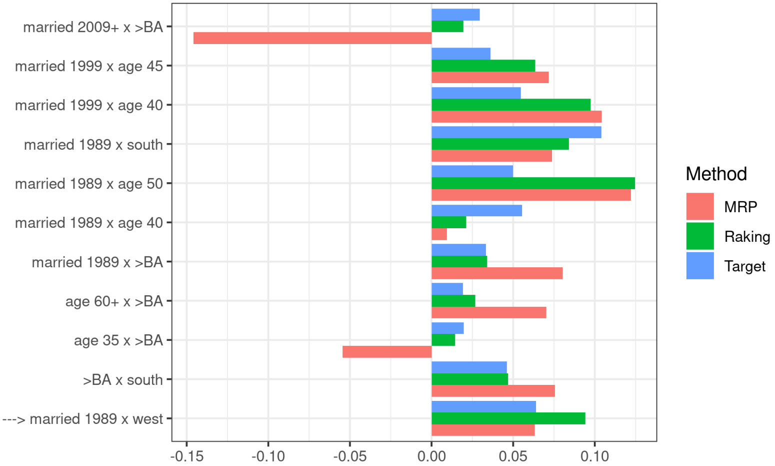


Figure 3: Imbalance plot for select interaction effects

Predictions

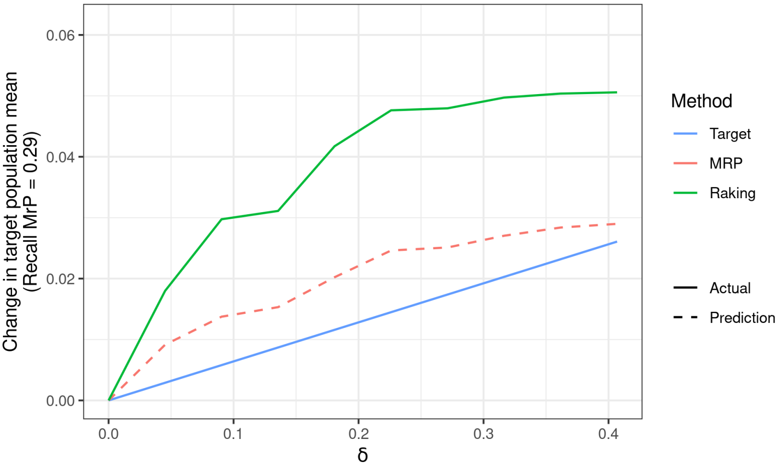


Figure 4: Predictions for the name change dataset

Predictions and actual MCMC results

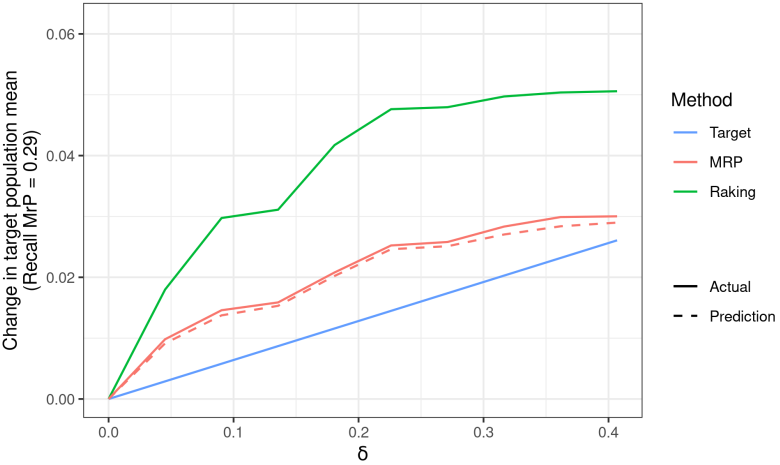


Figure 5: Predictions and refit for the name change dataset

Running ten MCMC refits: 28 hours Computing approximate weights: 27 seconds

Related and future work

Today, I focused on covariate balance. In this work, we also provide rigorous justification for

- Frequentist covariance estimation
- Parital pooling
- Negative weights (extrapolation)

Related and future work

Today, I focused on covariate balance. In this work, we also provide rigorous justification for

- Frequentist covariance estimation
- Parital pooling
- Negative weights (extrapolation)

Student contributions and future work:

- **Alice Cima** contributed significantly to this work
- **Vladimir Palmin** is working on extending MrPlew to `lme4`
- **Sequoia Andrade** is working on generalizing to other local sensitivity checks
- **Lucas Schwengber** is working on novel flow–based techniques for local sensitivity

Alice Cima

No picture!

Vladimir Palmin

Sequoia Andrade

Lucas Schwengber

References

Alexander, M. (2019). *Analyzing name changes after marriage using a non-representative survey*. URL: https://www.monicaalexander.com/posts/2019-08-07-mrp/.

B., Eli, Avi F., and Erin H. (2021). *Multilevel calibration weighting for survey data*. arXiv: 2102.09052 `[stat.ME]`.

Blue Rose Research (2024). *2024 Election Retrospective Presentation*. https://data.blueroseresearch.org/2024retro-download. Accessed on 2024-10-26.

Bonica, A. et al. (Apr. 2025). *Did Non-Voters Really Flip Republican in 2024? The Evidence Says No.* https://data4democracy.substack.com/p/did-non-voters-really-flip-republican.

Bürkner, Paul-Christian (2017). “brms: An R Package for Bayesian Multilevel Models Using Stan”. In: *Journal of Statistical Software* 80.1, pp. 1–28. DOI: 10.18637/jss.v080.i01.

Chattopadhyay, A. and J. Zubizarreta (2023). “On the implied weights of linear regression for causal inference”. In: *Biometrika* 110.3, pp. 615–629.

Cohen, P. (Apr. 2019). *Marital Name Change Survey*. DOI: 10.17605/OSF.IO/UZQDN. URL: osf.io/uzqdn.

G. and T. Broderick (2024). *The Bayesian Infinitesimal Jackknife for Variance*. arXiv: 2305.06466 `[stat.ME]`. URL: https://arxiv.org/abs/2305.06466.

G., T. Broderick, and M. I. Jordan (2018). “Covariances, robustness and variational bayes”. In: *Journal of machine learning research* 19.51.

Gelman, A. (2007a). “Rejoinder: Struggles with survey weighting and regression modelling”. In: *Statistical Science* 22.2, pp. 184–188.

— (2007b). “Struggles with survey weighting and regression modeling”. In.

Kasprzak, M., **G.**, and T. Broderick (2025). *How good is your Laplace approximation of the Bayesian posterior? Finite-sample computable error bounds for a variety of useful divergences*. arXiv: 2209.14992 `[math.ST]`. URL: https://arxiv.org/abs/2209.14992.

Ruggles, S. et al. (2024). *IPUMS USA: Version 15.0 [dataset]*. DOI: 10.18128/D010.V15.0. URL: https://usa.ipums.org.

Solon, G., S. Haider, and J. Wooldridge (2015). “What are we weighting for?” In: *Journal of Human resources* 50.2, pp. 301–316.